Community Issue Review: Crystallizing Knowledge for Encouraging Civic Engagement

Feng Sun

College of Information Sciences and Technology
Pennsylvania State University
University Park, PA 16802
fzs122@psu.edu

Guoray Cai

College of Information Sciences and Technology
Pennsylvania State University
University Park, PA 16802
cai@ist.psu.edu

ABSTRACT

Civic engagement in public decision-making requires that people form opinions or judgments based on the insights and knowledge around a public issue. Such knowledge is often not easily accessible by citizens because information about a policy issue tends to be scattered and buried in large and complex sources of potentially diluted and messy data. This creates information barriers that prevent ordinary citizens from effectively participating in government issues. We present an approach to addressing this problem that involves crystallizing the bits and pieces of knowledge scattered in the data into a form of knowledge that is well-framed, compact, and trustworthy to the citizens. This approach is articulated as a knowledge crystallization framework that specify the goals and processes of crystallizing policy knowledge. Following this framework, we present a concrete process, *community issue review* (CIR), that can be practiced as public decision analysis tools in a variety of community contexts. We implemented and used CIR in evaluating a real community proposal and observed the positive impact on the level of engagement. In the same time, we received feedback on how the CIR process and the supporting technology can be improved or be made more flexible to support best practices.

CCS CONCEPTS

•Applied computing → E-government; •Information systems
→ Collaborative and social computing systems and tools;

KEYWORDS

Knowledge crystallization, community issue review, civic engagement, e-participation

ACM Reference format:

Feng Sun and Guoray Cai. 2017. Community Issue Review: Crystallizing Knowledge for Encouraging Civic Engagement. In *Proceedings of dg.o '17, Staten Island, NY, USA, June 07-09, 2017*, 10 pages.
DOI: <http://dx.doi.org/10.1145/3085228.3085293>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

dg.o '17, Staten Island, NY, USA

© 2017 ACM. 978-1-4503-5317-5/17/06...\$15.00

DOI: <http://dx.doi.org/10.1145/3085228.3085293>

1 INTRODUCTION

Using information and communication technologies to improve the democratic decision-making process is one of the important but unfulfilled promises of digital government research [25, 35, 41]. Democracy empowers the public through their influence on political decisions [19]. Complex and contentious policy issues require not only a broad participation (by a larger part of the stakeholders), but deep participation (improved ability to deliberate and form considered opinions). However, meaningful participation in public decision-making is premised on the ability of the participants to produce reasonable, well-informed opinions [7, 11]. Such ability is often hindered by a lack of and selective exposure to a diverse marketplace of ideas [12, 37]. To date, digital government research has mainly focused on methods and infrastructures for collecting and sharing ideas, galvanizing interest and eliciting participation [25], but falls well short of addressing the complexity of participative government.

Online participation platforms have significantly enriched the channels of expressing opinions on public matters [17], but at the same time, it creates information glut to citizens' use. Typically, information about an issue is buried in online community forums, government web pages, documents, reports, newspaper articles, community meeting notes, as well as in the minds of experts, residents, and other stakeholders. Messages in these data are poorly framed and viewpoints are opinionated. Data about an issue are piece-wise, difficult to connect, redundant, and inconsistent. They may include false information, overly detailed information, or unverified (untrustworthy) information that creates more confusion than informing the public. Therefore, useful knowledge has to be discovered from the information and be contextualized for certain tasks [40].

Even if we assume that citizens are motivated to participate, it is widely known in the literature of deliberative democracy [23] and policy communication studies [39] that the public is facing tremendous information barrier to public participation. Elliman et al. [13] emphasized that the most fundamental barrier in public deliberations is the large amount of heterogeneous knowledge that needs to be made explicit in different formats at different stages. On one hand, average Americans have insufficient knowledge about political issue and government (known as "civic knowledge deficit") in order to practice direct democracy effectively [3, 6, 26]. On the other hand, public involvement processes furnish citizens with insufficient information about policy problems. When asked for opinions, citizens feel that they know little about the available choices among policy solutions, confused by the (often flawed) criteria for choosing among such solutions, and rushed to judgment before sufficient

reflection on those options [20]. To compensate this knowledge deficit, enthusiastic participants have no choice but having to deal with the daunting task of seeking, interpreting and synthesizing a large number of documents, comments, and claims, along with other materials to gain insight that informs opinions [22, 44]. Given that civic participation is only a small piece of individuals' life, it is not hard to imagine that majority of the public will not take on this challenge and they are likely to be excluded. Unfortunately, this problem has by far received little attention, and no practical solution has been proposed.

This paper argues for the need to communicate policy-relevant knowledge more effectively to the public by taking into account their limited cognitive capacity and attentional resources. This need can be met by incorporating an explicit phase of "knowledge crystallization" before engaging the broader public for their policy preferences. We present a conceptual framework for structuring knowledge crystallization tasks (section 3). Following this framework, we propose a concrete process, *community issue review* (CIR), that can be practiced as a policy-public communication tool in a variety of policy-decision contexts (section 4). We implemented and used CIR in evaluating a real community proposal and observed the positive impact of the level of engagement (section 5 and section 6). In the same time, we received feedback on how the CIR process and the supporting technology can be improved and be made more flexible to support best practices (section 7).

2 RELATED WORK

Informing the public with adequate knowledge about policy issues is hard for many reasons. For complex policy issues, a true understanding of a complex policy issue must be the result of synthesizing three types of knowledge. First, the public needs to be informed by good science that characterizes the potential consequences (benefits and risks) of any policy option [8]. Second, it must give all stakeholders an opportunity to express their social, economic, and ethical concerns. Third, it must be informed by an understanding of the institutional, political, legal, and operational contexts of decision-making. This knowledge can be buried in a plethora of online and offline information sources. With the advent of user generated content and social software systems, citizens have much broader spectrum of channels for public expressions, but in the same time, it has become increasingly difficult for interested citizens to identify and track all the appropriate information of interest and citizen comments or other responses to a given problem or set of issues. In a small town, like State College Borough, information on a policy issue has become widely scattered across multiple websites (e.g., government, schools, community groups, news organizations) as well as web-based user-generated content (e.g., posts on public Facebook pages, individual blogs, Twitter messages or image collections).

To address the problem of the wide dispersion of local knowledge, various computer-mediated systems and data mining techniques have been developed to automatically discover and aggregate diverse sources. Kavanaugh et al. [27] developed a *Virtual Town Square* (VTS), a local news aggregator that affords civic interaction through tagging, commenting, and sharing insights. However, even data are aggregated, they may still be too large for policy

makers to make sense of them. These harvested news can be further analyzed to understand trends and events in social contexts. Automated textual analysis tools have been used to detect important messages and alert analysts. Another solution proposed in a rule-making context [16] is to provide information shortcuts for citizens that can enable them to behave civically smarter, but it also raised concern that designers have the power to frame citizens' understanding of the information. Hagen et al. [24] automatically analyzed thousands of petitions to generate more concise reports for decision makers. Topic modeling methods [45] are useful here because it summarizes the most popular topics that appear news articles and blogs and representing them together in an intuitive way. Nevertheless, topics that are generated by topic modeling methods do not always represent the original contents well, since the best representation is context-dependent. Meaningful summary of large text data normally requires a manual inspection of the LDA output by domain experts to make sense of the analytical results.

While online forums and news aggregators help create a more effective public sphere for free flow and exchange of public information and political ideas, they do not address two challenges from the consumption side of these information resources. First, it does not take away the cognitive load in interpreting and synthesizing messages in the data, which will turn many potential participants away. Second, when people have very limited attentional resources allocated to information seeking from a vast pool of data, their attention will be selectively allocated to those items that are aligned with preexisting beliefs [38]. This is a well-known effect called "confirmation bias" [33]. Such selective exposure is likely to increase polarized attitudes on issues, impairing effective democracy [43].

Since our faith on the public competence of gathering policy knowledge has been lost, researchers have turned into methods that can delegate such task to a small group of citizens who will engage detailed investigation, analysis, and deliberation on an issue and report their findings back to the broader public. This Along this line, recent deliberative innovations such as *mini-publics* [21] and *random assembly* [20] are relevant. Mini-publics fit well for this purpose because they attempt to achieve impartiality, preference change and public reasoning by bringing together a random selection of non-partisan citizens to discuss key issues [14]. Mini-publics are made up of ordinary, non-partisan, lay citizens and are "designed to be groups small enough to be genuinely deliberative and representative enough to be genuinely democratic" [21]. In mini-publics, either a random or a stratified sample of the population is selected to achieve a microcosm of the population, with each citizen having an equal chance of being selected [30]. They are usually asked to deliberate on the issue given evidence and advocacy provided by subject matter experts [14]. However, past use of mini-publics has been mostly for the purpose of forming policy recommendations to decision-makers, rather than crystallizing knowledge for the public to be informed. The only exception is the Oregon's model of *Citizens' Initiative Review* (OCIR) [18]. OCIR randomly selects a panel of voters to evaluate a ballot measure through reviewing the information gained from campaigns and policy experts. The panel will finally generate a set of statements conveying the most important findings of the measure, which serves as an easily accessible resource for informing voters.

OCIR was designed to be an institutional tool for the state-level vote on ballot measures, and the practice of OCIR involves a face-to-face meeting that lasts 4-5 days, which may exclude certain population to participate. In contrast, our work focuses on local government policy issues and we attempt to reduce the participation barrier by adopting a hybrid (face-to-face and online) approach.

3 KNOWLEDGE CRYSTALLIZATION FOR CIVIC ENGAGEMENT

It is widely recognized that capturing all the relevant information and making it easily accessible to local citizens is an important step to facilitate civic engagement [11]. Providing people with access to more information is not the enough. Digital government technologies must go beyond a information repository, and provide help to citizens who can maximize their allocation of attention to information that will be useful for their judgment and opinion formation.

Knowledge crystallization is a process that aims at finding the most concise and compact description form of knowledge possible that can be discovered from a set of data relative to a given task. In the context of civic engagement, knowledge crystallization takes all the data that we can collect about a particular issue or subject, and puts them through a systematic process of distilling relevant nuggets, purifying, abstracting, and compacting to create a best and most accessible form of knowledge that can be understood and trusted by the public.

Crystallization is a metaphor borrowed from chemical engineering, where the goal of crystallization is to produce a highly purified and ordered crystal lattice from raw materials through the processes of purification and condensation [1]. Similarly, we define knowledge crystallization as a process that aims to produce a most insightful and compact description of the relevant knowledge from a data set without removing crucial information for a given task. Examples of knowledge crystallization tasks include writing a business intelligence newsletter, reporting on the analysis of a business strategic management practice, or a scientist writing a literature review article [4].

An overview of the knowledge crystallization process is shown in Figure 1. The rectangular boxes represent entities involved in the process. The arrows represent flow relationships among them. This process has four small loops and has one set of loops that cycle around knowledge evaporation and another that cycles around knowledge condensation, with plenty of interaction between these. A complex task is decomposed into a chained sequence of subtasks, where the output of one becomes the input to the next. This process proceeds guided by a schema and helps refine the schema. The processes and data are arranged by the amount of effort and degree of knowledge crystallization. A bigger rectangular task wraps the entire process and serves as the context. The process flow shows the transformation of information as it flows from raw data to reportable and accessible results under the guidance of a knowledge crystallization schema given a task.

We call this type of knowledge crystallization *Issue-based Knowledge Crystallization* (IBKC) since it is a special version developed for community issues. It is consistent with the work of Card et al. [5] who recognized that effective knowledge crystallization requires

Figure 1: The framework for issue-based knowledge crystallization

several subtasks to be performed: information foraging, knowledge schematization (searching for schema and representing findings using the schema), knowledge compaction, and knowledge communication. IBKC goes beyond the previous conception of knowledge crystallization by Pirolli and Card [5, 36] in two significant aspects:

- IBKC extends knowledge crystallization from individual to a group activity. As an ill-structured problem, crystallizing knowledge from a large volume of data can benefit from group work by a divide-and-conquer approach, complementary expertise and perspectives, and enriched judgments.
- IBKC specializes the knowledge crystallization problem for analyzing issue-based community knowledge to inform public deliberation. In the practice of democracy, the perception of “good information” is not universal [10]. Democratic deliberation requires that public opinion should be informed by a full and balanced understanding of a community issue and associated solutions. That imposes a strong constraint that the outcome of IBKC cannot lose any essential knowledge that is in the data.

4 COMMUNITY ISSUE REVIEW

Community Issue Review (CIR) is a community-level panel-based deliberation process for crystallizing knowledge about a pending community issue [29] that is specially tailored to the need of informing the public on local policy issues.

CIR is an implementation of issue-based knowledge crystallization framework we proposed earlier. In order for a CIR process to happen, three conditions must exist: (1) there is public issue that is pending for decision and is drawing the attention from the public, (2) there has been sufficient expressions of concerns, policy choices and preferences openly exchanged among the public, experts, and policy advisors, and (3) all the information sources of about a policy issue have been identified and collected into a collection of documents. The goal of CIR is to produce a set of *Citizens' Statements*

that answers all the questions that a voting community member wants to know [2]. In the same time, the *Citizens' Statements* must be short, concise and easy comprehension to minimize the time and effort required from citizens, and must be trusted by the public to be a non-partisan analysis based on the full consideration of public goods. In particular, CIR re-interpret the information provided by experts and translate them into easily accessible expressions without technical jargons, unnecessarily sophisticated analytical details. CIR uses a citizen panel as issue analysts who are likely to evaluate and deliberate on the issue the same way as their peer citizens, which can increase trust and acceptability of the statements presented to the broader citizens. The products generated through CIR can provide the community with insights concerning the issue so that everyone in the community is able to form opinions effectively and thus has an impact on the public decision-making process.

In the rest of the section, we describe CIR process in terms of phases and relevant activities as an instance of IBKC. We will also identify the challenges users may face and show how technical support can contribute.

4.1 Preparation

One important design question is: who should be the part of the CIR panel for crystallizing the community knowledge? This process should not be done by an individual and should not be done by government officials or experts. Instead, such knowledge crystallization is preferred to be done by a small group of citizens to ensure both in-depth analysis and democratic assessment. We follow the work of a small group deliberative democracy process [20], which selects panelists from the relevant population as representatives. Stratified random sampling is adopted in CIR to make sure that everyone affected by the issue has equal chance of being selected while different subgroups in the broader population are adequately represented.

In addition, the ideal input of CIR is all the information that can be found about a community issue, for example, proposals, ordinances, reports, and news articles. The subject matter experts and government officials will help organize and structure the information into a collection of documents. The contents usually include neutral descriptions of proposals/ordinances and evidence supporting/against an issue, and panelists are allowed to add more materials later if necessary.

Before moving to the first phase, the panelists will get together and learn about the community issue through a briefing. They will have a chance to get to know each other as collaborators. Tutorials for the online system will be provided that enables them to work on it during the four phases before the final-day meeting. They can also communicate with experts directly to resolve quick questions. Last but not least, as the panelists have a basic understanding of the community issue, they will be asked to come up with a knowledge scheme used for the following phases.

4.2 Phase I: Extract Information Nugget

Nugget Extraction in CIR is aimed at reliably recognizing and collecting all nuggets relevant to the pending issue, and it is the prerequisite for subsequent tasks of knowledge crystal formation, refinement, and compaction. The task involves a number of cognitive

and physical actions. First, the CIR panel members are charged with investigating a policy issue that is usually complex and controversial. In the same time, they are provided with a document collection that is considered to have all the data we can find about the issue. Other than published reports, websites, and news articles, the document collection also contains interviews with subject matter experts and their written statements. During this first phase of CIR, a citizen panel gathers information nuggets relevant to a policy issue through an online analytic forum where individuals can access all the documents to be analyzed, extract nuggets, and tag them by a particular theme. The collected nuggets are expected to cover all the information that can be found.

4.3 Phase II: Assemble and Improve Claims

The purpose of *Claim Assembly* in CIR is to transform collected information nuggets into claims, which should be relatively well-written, self-contained, and based on facts and evidence. There are two kinds of claims: findings (objective facts) and opinions (facts with implicit position). The opinions can be further decomposed into two categories: substantiate and refute, depending on the position.

The initialization of a claim is required to be from one or more information nuggets. Claims can be further elaborated and improved through adding more information nuggets, removing irrelevant or unimportant contents, rewording, and paraphrasing, merging several claims or splitting a claim. This phase requires the panel members to analyze the collected information nuggets carefully, either individually or collaboratively, and compile them into a better format.

4.4 Phase III: Generate Statements

Claims might still be reduplicated, conflicting and redundant and thus cannot meet the requirements of the compactness. This phase improves the qualities of claims and makes them more compact, defensible, and understandable. Once claims are of high quality that meets the requirements of the task, they are considered to be exquisite statements.

Several operations can be done upon such claims, such as rewording, merging, splitting. A group of people sort out the essence and discard the dross by performing aforementioned operations, as well as discussing, categorizing (categorize a claim as substantiate, refute or finding), prioritizing and voting. Scheme plays an important role here by ensuring that each dimension of the issue is fully reviewed and accepted by the majority.

4.5 Phase IV: Communicate Statements

In this phase, the claims are compiled into a set of statements of manageable length. These statements must be presented in a way that is easily accessible and available to local citizens, enabling them to acquire necessary issue-relevant knowledge to participate. For example, the vocabulary and sentence structure used by the statements have to be "translated" from the languages that only experts and professionals understand to be a form less likely to overwhelm and usable by the general public.

The final statement contains ten statements that summarize the issue and why it is important to the community, five statements are

Figure 2: The overview of community issue review as a knowledge crystallization implementation

the strongest arguments in favor of the issue and five statements are the strongest arguments against the issue. Considering the audiences are the general public, technologies such as visualization and storytelling, could be adopted to make statements more accessible, interactive, and easy to understand.

4.6 Facilitative Moderation

Civic participation nowadays requires expert facilitators to be integrated into the system to better support democratic deliberation and enlightened understanding [28]. Firstly, conflicts among views on issue-relevant information with different understandings, values, and knowledge are almost certain and must be properly managed [32]. Secondly, different levels of computer skills and familiarity with issues may vary greatly. Thirdly, there exists a variety of contribution patterns that cannot be captured by the system fully. Even with instructions and training sessions that explain how to participate effectively, many participants struggle with making statements by providing evidence, reasons, and factual support. In addition, facilitators could also coordinate with the panelists and experts in terms of information exchange and take care of all other things irrelevant to the issue. In short, facilitation is an essential component to ensure the effective implementation of CIR.

5 IMPLEMENTATION

We implemented CIR as an instance of knowledge crystallization in a hybrid mode that includes both face-to-face meetings and online interactions. This section describes the technical supports provided by the online system to address the difficulties panelists may encounter.

5.1 Combining Face-to-face and Online Interaction

CIR can be conducted in a purely face to face environment. However, on the one hand, citizens have their daily work and can only allocate limited time (especially daytime) and effort for CIR. Besides,

information technologies provide many features that can enhance the cognitive capability of people. Therefore, we engineered CIR as a blending of online and face-to-face activities that lasts about ten days to two weeks depending on the complexity of tasks. There are two face-to-face meetings on the first and last days respectively, and the panel works online during the time in between.

There are a lot of differences between face-to-face and online interactions, especially in the context of deliberation. For example, a lack of computer skills may limit a person's capability of participating, and the possibility of asynchronous communications will also pose different collaboration patterns. Many of the consequences still require further investigations.

5.2 Supporting Nugget Extraction

Figure 3 shows the user interface of the system designed to support Phase I of CIR. The interfaces for all the four phases follow a similar two-column design. The left column in such design corresponds to the input of the current phase while the right one is the expected output. In the Nugget Extraction Phase, the left column is a *document view* and a list of information nuggets is placed in the right column.

The document view contains a collection of documents. A table of contents button allows panelists to navigate through all documents. Panelists can select one document and start to extract nuggets by selecting a piece of text judged as relevant to one or more theme. Once a segment of text is selected, it will be highlighted with yellow color and invoke a pop-up menu. Before submitting an extraction, there is one more step that asks panelists to assign this information nugget to a theme. All the themes are listed on the top and a detailed explanation will be provided when scrolling over each theme icon with the mouse. Highlighted contents are coded in blue in the document view.

All the extracted nuggets are collected into the *nugget list*. Once a nugget is newly extracted, it will immediately appear in the top of the list. In case panelists assign a nugget to an incorrect theme, one can modify the assignment by the "reassign" action that is

Figure 3: The interface for nugget extraction

available from the context menu beneath a nugget. A nugget can also be removed if it is a duplicate or a mistake. Nugget List view is actively linked to the Document View, allowing panelists to trace back to where a nugget originates in the document. Capturing the relationship between nugget and its origin in documents effectively makes it possible to replay and review the analytical process later on [34].

5.3 Supporting Claim Making

Figure 4 shows the user interface of the system designed to support Phase II of CIR. *Nugget list* is placed on the left side and *claim space* is placed on the right side.

Figure 4: The interface for claim making

Claim space is a combination of claim list and claim workspace. By default, the claim list is shown on the right side. Claim list, as its name shows, contains a list of claims. A new claim can be created by clicking either the “adopt” button that associates with each entry in the nugget list or the “plus” button on the top of the claim list. Once a new claim is created, the view automatically switches to the claim workspace where one can work on a claim and discuss it. The workspace of an existing claim can be evoked by clicking its entry in the claim list. Panelists can switch back to the full list of claims with the “back” button.

While working in the claim workspace, since a claim has to be derived from one or more information nuggets, panelists are expected to select some entries from the nugget list. By adopting them, their contents are copied into the claim editor, and the links between the selected nuggets and the current claim are built. From there, panelists can further elaborate on the texts to make them relatively well-written and self-contained. In addition to editing the claim directly, panelists are also encouraged to make comments and questions on the claims. They are expected to collaboratively create and improve claims. All the relevant activities, including the adding and removing of nuggets, the modifications of claims, and the discussions are logged and shown in the bottom of the claim workspace.

Nugget list in this interface shares most of the functionalities available in the previous phase, except that an additional “adopt” button is provided for each nugget entry. On the one hand, as described above, it enables panelists to initiate a claim based on the associated nugget. On the other hand, if a nugget has been used, the number of its usage shows, which enables panelists to retrieve all the claims that adopt the nugget and allows them to be aware of how the nuggets are used in these claims.

5.4 Supporting Statement Composition

The composition of statements is implemented as two steps in the online platform: categorization and refinement. Categorization is aimed to determined whether a claim is a finding, or it substantiates or refutes the community issue. The goal of refinement is to compact claims while maintaining their informativeness and readability.

Figure 5: The interface for categorization

The user interface of categorization is shown in Figure 5. The task in this step is to assign the claims listed on the left side to the corresponding category located on the right side. The expected outcome of CIR is a set of Citizens' Statements, including 10 findings, 5 pros, and 5 cons. The slots associated with the 20 statements are allocated in advance and can be assigned with customized labels for easier reference. Each category is equipped with a progress bar that shows the current working process of categorization.

The categorization is performed by dragging and dropping claims from claim list to a category slot. A statement is composed by referring to one or more claims. A claim can also be used in multiple

statements, even multiple categories. For example, a finding, with different interpretations, can serve as a pro and/or a con. This many-to-many relationship is captured and shown in the interface, which allows panelists to be aware of how each claim is categorized and how each slot is composed.

Figure 6: The interface for refinement

The user interface of refinement is shown in Figure 6. This step is intended to compact claims that contribute to a similar topic and produce a well-written and informative statement for each slot. By clicking a slot, the corresponding workspace will show. This workspace allows panelists to produce statements and improve on them. There might exist multiple versions of a statement suggested by panelists but only one version can be adopted at one time. For each slot, panelists are expected to discuss its multiple variants, refine the contents, and produce a commonly agreed statement.

5.5 Supporting Statement Presentation

The finalized statements are structured using a set of *Frequently-Asked-Questions* (FAQ) (see Figure 7). This presentation style enables the general public to find answers to their questions quickly. When the general public wants to seek more details about any part of the Citizens' statements, they can follow the hyperlinks to trace back to the origins of those ideas, leave comments, or engage conversations with other users.

Figure 7: The interface for statement presentation

For each statement, links are provided that lead to the original information nuggets, allowing panelists to find out the context of the statement, as well as how these information nuggets are transformed into the current statement. Also, for some unfamiliar and unintelligible terms and phrases, definitions and explanations are provided through a mouse-over glossary. References to additional supplementary materials from external sources are provided if necessary.

5.6 Supporting Collaboration

In addition to entity-eccentric discussion and editing, collaboration among panelists is supported consensus building through a chat room. Communications among panelists and subject matter experts are supported through a question panel.

Figure 8: Chat room (left) and question panel (right)

A *chat room* is available to panelists in all phases that allows them to communicate spontaneously, ask questions, and leave notes for each other. Asynchronously communication is also supported as all the chatting messages are stored. Discussion around a particular entity, e.g., a claim, is expected to be conducted in the discussion board associated with the entity. While the chat room is designed for communicating things that don't fit anywhere else and coordinating group work.

Panelists can directly ask a question related to a community issue in the question panel. Normally, the question panel is minimized but it can be evoked at anytime by clicking on the yellow question mark positioned on the upper right corner. Some of the questions can be answered immediately by peers. When questions can not be resolved within panelists, they are forwarded to subject matter experts. Once answers are received from experts, they will be incorporated into the document collections.

5.7 Supporting Facilitation

All the phases of CIR, including face-to-face meetings, are facilitated by a professional moderator to ensure the quality of the collaboration and work.

Figure 9: Facilitators' dashboard

The facilitator is provided with a control panel in the online system, which currently consists components of an overview of panelists' activities, phase control, scheme management, and document management. The overview is a crafter visual representation of panelists that helps facilitators to monitor the situation and making decisions about facilitation strategies. The phase control allows facilitators to switch phases. Facilitators are enabled to edit schemes and documents as well. Although the functions are still limited, as facilitation strategy is really critical to the performance of CIR, more flexible and efficient means should be considered as we learn more about the process.

6 CASE STUDY

In order to test our implementation of CIR and iteratively improve it, we conducted a case study to gather the feedback from the potential panelists. We are particularly interested in answering the following questions: What are good knowledge crystallization scheme? How should the contents (e.g., documents, extracted nuggets) be organized and presented? How to define and implement various phases to better structure the process? What is the usability of the online system and how it can be improved? How do participants collaborate in CIR?

6.1 Procedure

The community issue used in this study is *inflationary tax indexing*. There are four ways that the State College (Pennsylvania) Borough generates revenue that is consistent and regular revenue: real estate tax, earned income tax, local services tax, and real estate transfer tax. Among them, real estate tax is the major source. However, the base upon which the real estate tax is levied has not changed significantly over the past 10 years. There isn't much room for Borough residents to continue to develop in the Borough when the natural base is stagnant. But the costs of the services provided within the Borough continue to grow at least with inflation and in some cases more than inflation. So the proposal is a recommendation to council that real estate tax should be increased by at least inflation every year just to keep pace with the cost of providing services to the Borough.

Considering that the issue is closely related to real estate property and State College is a college town, renters (most of them

are students) and homeowners should be involved. In fact, we recruited fourteen participants as the citizen panel. Most of them were recruited via mailings that were sent at random based on the addresses provided the borough office. Three of them from specific student organizations were recruited via targeted email. The three students rent in the borough. The rest are homeowners. We started with fourteen panelists and ten of them attended the final meeting. There are also four subject matter experts involved. Two of them are proponents that support the inflationary tax indexing, and two are opponents against the issue.

The study was initially designed to follow the typical mode: ten days that consists of a face-to-face introduction meeting on the first day, another face-to-face finalization meeting at the end, the four major phases of CIR are conducted in the online environment. However, the study did not proceed as expected and thus we added three additional work sessions and one more face-to-face meeting to reflect on the implementation of CIR.

Figure 10: Face-to-face meeting on Day 1

On the first day, a short introduction and a training session were conducted in the face-to-face meeting. In the first session, the instructor introduced the community issue review in terms of its process, expected outcome and various roles. Then the community issue of inflationary tax indexing was introduced, followed by a question and answer period that allows the panelists to ask subject experts to answer questions and address concerns. The panelists are then asked to come up with a set of themes that help organize the process.

In the training session, the instructor firstly explained to the participants the user interface and available functions of the online system. After that, the participants were given about 1 hour to interact with the system using sample data with the help of the instructor. One of the authors served as technical support to the panelists to resolve any problems they had.

After the meeting, the panelists went home and began to work in the online environment. They were expected to follow CIR phases and collaboratively produce candidate statements for the final-day meeting to discuss. During this period, a facilitator kept monitoring the online activities and moderate the process when necessary. Instructional emails were sent every morning, along with a summary of the progress by far. A video tutorial of the online system is also available online.

Nevertheless, at the moment of final meeting, due to the task complexity, the participants were close to finishing claim making phase and had not started statement composition phase yet. The problems we identified were mainly about the usability of the system and the computer skills of the participants. Therefore, in the final-day meeting, we organized another tutorial session in which our research team members provided one-on-one instructions of the system.

We further extended this study with three more work sessions and another meeting. In total five participants attended the three work sessions which led to a set of candidate statements. In the last meeting, they discussed these candidates deliberatively and finalized the citizens' statements.

As this is still a preliminary study and we wanted to gather input from the participants as much as possible, an exploratory approach was employed to collect their feedback. During each work session, they described what they were doing and thinking aloud, especially the intentions behind behaviors. We only provided guidance and assistance whenever the participants had confusion.

6.2 Evaluation and Feedback

Based on the participants' feedback and our observations, we compiled a series of points and summarized them in this section.

6.2.1 Knowledge Crystallization Scheme. In this study, we adopted a set of themes as the knowledge crystallization scheme. Each theme represents one important aspect of the issue and is shown as a phrase with detailed explanations on demand. For example, "Affordability" was used in the study as a theme, which indicates how a tax increase affects the price of owning and renting properties.

All the participants reported that the themes were only useful for the first phase to organize extracted nuggets. It was difficult for them to continue to use the themes as the extracted nuggets were transformed into claims. Instead, some of the participants suggested that several guiding questions would be more helpful for all phases.

6.2.2 Phase Transition. We thought it is better to organize CIR as an explicitly phase-based process following the IBKC framework for it being more manageable. Phases switch only when all participants feel that the time is right for evaluating and using the found information the input for the next phase. By enforcing phase-based process, panelists were expected to work synchronously and thus their contributions could be evaluated and utilized collaboratively.

However, Some participants reported that it was difficult for them to divide phase clearly; it caused confusion to them as they had to understand exactly the design of each phase. To address this problem, we enhanced our system by allowing panelists to be able to do all the work in one integrated interface where phases are implicitly enforced. We presented the revised interface to two participants and received positive feedback.

6.2.3 Learnability and Accessibility. The most common issue pointed by the participants was the usability of the online system. The participants were unaware of many available features and sometimes used the system in an incorrect way. The targeted users are ordinary citizens, among which many have insufficient computer skills. Therefore, on the one hand, the system should be

designed to be easier to learn and operate. On the other hand, a more sophisticated technical support and training session should be provided.

6.2.4 Collaboration, Coordination, and Communication. Our observations show that participants communicated a lot in face-to-face meetings while they worked almost individually in the online environment, though a variety of communication channels were provided. One participant believed it is due to time delays in asynchronous communication while people do expect immediate responses or in-time notification. This was explained by [9].

One improvement is to provide subscription/notification service: Once a participant makes a contribution to an entity, she is considered to subscribe to the related thread. Whenever there is an update, e.g., another participant leaves a comment, the participant will be notified. The idea of the private and public workspace was also mentioned by some participants, which allows them to work in their private workspace and share with others only when necessary.

6.2.5 Flexibility of Organizing and Retrieving Contents. Currently all the entities involved in CIR, such as documents, extracted nuggets, assembled claims, and candidate statements, are structured in a linear fashion. Although some filters are provided that allow panelists to select, for example, the information nuggets tagged by a particular theme, it is still limiting the way of organizing them. Some participants would like to see the system provide more means of structuring the entities, for example, to cluster documents based on contents in advance.

Another feature the participants hope to have is a search function. We intentionally removed the search function as a way to enforce people to go through all the documents thoroughly rather than doing a keyword search when doing sensemaking. However, the participants do have a need to revisit what they have read, and search function can support that. Bookmarking is also a solution to that.

7 DISCUSSION

Issue-based knowledge crystallization is empowered by a mini-public. Currently, Mini-publics are widely used in deliberative events, such as citizens' juries, consensus conference, planning cells, deliberative polls, citizen assemblies [15]. Mini-publics contribute to these decision-making processes by providing collective recommendations and opinions through a sample of citizens. Instead of making decisions directly, in CIR, the mini public function as an information processor and is used to analyze the documents without expressing any opinions. Ideally, the outcome should cover all important information related to an issue objectively that is adequate for the broader citizens to form opinions.

In addition, online participation plays an important role in CIR. It offers the flexibility of time and place by providing citizens with more channels of participating in deliberative democracy process, but also introduces the limitation of generalizability because of different computer skills [31]. In our study, college students possess better computer skills than the general public, and this might have influenced the outcome of online participation. Our implementation of CIR addresses such limitation through a training session on the first-day meeting and technique support throughout the process,

but it still requires panelists to have basic computer skills, which an increasing number of the general public is acquiring.

8 CONCLUSION AND FUTURE WORK

In this paper, we identified the challenges caused by information overload and knowledge deficit that prevent ordinary citizens from participating public life effectively. Drawing from observations of how local government decision-making works and theories of information, we proposed CIR, an instance of knowledge crystallization, as a solution to the above problem. We implemented CIR with online technologies and presented it to a group of panelists. Based on lessons learned from their feedback, we gain a better understanding of CIR and are aware that there is a need for future research that continues to improve this process.

The next step, from the technical perspective, in addition to the features required by the participants of the case study, we plan to incorporate visualization features that is more intuitive for citizens to visually interact with the contents in the system. Some effort has been made for supporting nugget extraction phase visually [42]. Preprocessing of information will also be useful to enhance people's analytical capability. From the people side, we are also interested in the human behaviors during CIR, especially the collaborative and deliberation patterns in terms of how they influence the quality of CIR outcomes.

ACKNOWLEDGMENTS

The authors would like to acknowledge funding support from National Science Foundation under award # IIS-1211059.

REFERENCES

- [1] Wolfgang Beckmann. 2013. *Crystallization: Basic Concepts and Industrial Applications*. Wiley-VCH, 362 pages.
- [2] Abdelhamid Boudjelida and Sehl Mellouli. 2016. A Multidimensional Analysis Approach For Electronic Citizens Participation. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research - dg.o '16*. ACM Press, New York, New York, USA, 49–57.
- [3] Stephanie Burkhalter, John Gastil, and Todd Kelshaw. 2002. A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Communication Theory* 12, 4 (2002), 398–422.
- [4] Stuart K Card, Jock D Mackinlay, and Ben Shneiderman. 1999. *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- [5] S K Card, J D Mackinlay, and B Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers, San Francisco.
- [6] Michael X Delli Carpini and Scott Keeter. 1997. *What Americans know about politics and why it matters*. Yale University Press.
- [7] Simone Chambers. 2003. Deliberative Democratic Theory. *Annual Review of Political Science* 6 (2003), 307–326.
- [8] National Research Council. . Washington, D.C.
- [9] David D Curtis and Michael J Lawson. 2001. Exploring collaborative online learning. *Journal of Asynchronous learning networks* 5, 1 (2001), 21–34.
- [10] Brenda Dervin. 1994. Information-Democracy: An Examination of Underlying Assumptions. *Journal of the American Society for Information Science*. 45, 6 (1994), 369–385.
- [11] John Dewey. 2012. *The public and its problems: an essay in political inquiry*. Pennsylvania State University Press.
- [12] James N Druckman, Jordan Fein, and Thomas J Leeper. 2012. A source of bias in public opinion stability. *American Political Science Review* 106, 02 (2012), 430–454.
- [13] T Elliman, A Macintosh, and Z Irani. 2007. A model Building Tool to support Group Deliberation (eDelib). *International Journal of Cases on Electronic Commerce* 3, 3 (2007), 33–44.
- [14] Stephen Elstub. 2014. Mini-publics: issues and cases. In *Deliberative Democracy: Issues and Cases*, Stephen Elstub and Peter McLaverty (Eds.). Edinburgh University Press.
- [15] Stephen Elstub and Peter McLaverty. 2014. *Deliberative democracy: Issues and cases*. Edinburgh University Press.
- [16] Cynthia R Farina, Mary J Newhart, Josiah Heidt, and Jackeline Solivan. 2013. Balancing Inclusion and "Enlightened Understanding" in Designing Online Civic Participation Systems: Experiences from Regulation Room. *Proceedings of the 14th Annual International Conference on Digital Government Research* (2013), 180–189.
- [17] Dennis Friess and Christiane Eilders. 2015. A systematic review of online Deliberation Research. *Policy and Internet* 7, 3 (2015), 319–339.
- [18] John Gastil and Katie Knobloch. 2010. Evaluation report to the Oregon state legislature on the 2010 Oregon Citizensfi Initiative Review. *Unpublished report*, University of Washington, Seattle, WA. (2010).
- [19] J Gastil and P Levine. 2005. *The Deliberative Democracy Handbook: Strategies for effective civic engagement in the 21st century*. Jossey-Bass, San Francisco, California.
- [20] John Gastil and Robert Richards. 2013. Making Direct Democracy Deliberative through Random Assemblies. *Politics & Society* 41, 2 (2013), 253–281.
- [21] Robert E. Goodin and John S. Dryzek. 2006. Deliberative Impacts: The Macro-Political Uptake of Mini-Publics. *Politics & Society* 34, 2 (2006), 219–244.
- [22] Bertram Myron Gross. 1964. *The managing of organizations: The administrative struggle*. Vol. 2. [New York]: Free Press of Glencoe.
- [23] Niklas Gudowsky and Ulrike Bechtold. 2013. The Role of Information in Public Participation. *Journal of Public Deliberation* 9, 1 (2013).
- [24] Loni Hagen, Teresa M Harrison, Özlem Uzuner, Tim Fake, Dan Lamanna, and Christopher Kotfila. 2015. Introducing Textual Analysis Tools for Policy Informatics: A Case Study of e-Petitions. In *Proceedings of the 16th Annual International Conference on Digital Government Research (dg.o 2015)*. 10–19.
- [25] Janette Hartz-Karp and Brian Sullivan. 2014. The unfulfilled promise of online deliberation. *Journal of Public Deliberation* 10, 1 (2014), 1–5.
- [26] Lawrence R Jacobs, Fay Lomax Cook, and Michael X Delli Carpini. 2009. *Talking together: Public deliberation and political participation in America*. University of Chicago Press, Chicago.
- [27] Andrea Kavanaugh, Ankit Ahuja, Manuel Pérez-Quinones, John Tedesco, and Kumbirai Madondo. 2013. Encouraging Civic Participation Through Local News Aggregation. *Proceedings of the 14th Annual International Conference on Digital Government Research* (2013), 172–179.
- [28] Ian Kearns, Jamie Bend, and Beatrice Stern. 2002. *E-participation in local government*. Institute for Public Policy Research.
- [29] Jessica N Kropczynski, Guoray Cai, and John M Carroll. 2015. Understanding the Roles of Artifacts in Democratic Deliberation from the Citizensfi Initiative Review. *Journal of Social Media for Organizations* 3, 2 (2015), 1–22.
- [30] Cristina Lafont. 2015. Deliberation, Participation, and Democratic Legitimacy: Should Deliberative Mini-publics Shape Public Policy? *Journal of Political Philosophy* 23, 1 (2015), 40–63.
- [31] Seong-Jae Min. 2007. Online vs. face-to-face deliberation: Effects on civic engagement. *Journal of Computer-Mediated Communication* 12, 4 (2007), 1369–1387.
- [32] Christopher W Moore. 2014. *The mediation process: Practical strategies for resolving conflict*. John Wiley & Sons.
- [33] Raymond S Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology* 2, 2 (1998), 175.
- [34] Chris North, Remco Chang, Alex Endert, Wenwen Dou, Richard May, Bill Pike, and Glenn Fink. 2011. Analytic Provenance: Process+Interaction+Insight. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 33–36.
- [35] Panagiotis Panagiotopoulos, George Gionis, John Psarras, and Dimitris Askounis. 2011. Supporting public decision making in policy deliberations: an ontological approach. *Operational Research* 11, 3 (2011), 281–298.
- [36] P Pirolli and S Card. 1999. Information foraging. *Psychological Review* 106, 4 (1999), 643–675.
- [37] Vincent Price and Peter Neijens. 1997. Opinion Quality in Public Opinion Research. *International Journal of Public Opinion Research* 9, 4 (1997), 336–360.
- [38] Eric Rassin, Anita Eerland, and Ilse Kuijpers. 2010. Let's find the evidence: An analogue study of confirmation bias in criminal investigations. *Journal of Investigative Psychology and Offender Profiling* 7, 3 (2010), 231–246.
- [39] Garrett Richards, Kenneth Belcher, and Bram Noble. 2013. Informational Barriers to Effective Policy-Public Communication: A Case Study of Wind Energy Planning in Saskatchewan, Canada. *Canadian Public Policy* 39, 3 (2013), 431–450.
- [40] Jennifer Rowley. 2007. The Wisdom Hierarchy: Representations of the DIKW Hierarchy. *Journal of Information Science* 33, 2 (April 2007), 163–180.
- [41] Hans J Schnoll. 2015. *E-government: Information, technology, and transformation*. Routledge.
- [42] Feng Sun, Ye Tian, and Guoray Cai. 2016. Supporting Collaborative Information Seeking in Online Community Engagement. In *Collaboration Technologies and Systems (CTS), 2016 International Conference on*. IEEE, 241–248.
- [43] Cass R Sunstein. 2001. *Echo Chambers: Bush v. Gore, Impeachment, and Beyond*. Princeton University Press Princeton.
- [44] Alvin Toffler. 1990. *Future shock*. Bantam.
- [45] Zhijun Yin, Liangliang Cao, Quanquan Gu, and Jiawei Han. 2012. Latent Community Topic Analysis: Integration of Community Discovery with Topic Modeling. *ACM Transactions on Intelligent Systems and Technology (TIST)* 3, 4 (2012), 1–21.